Measurement error in earnings data in the Health and Retirement Study

Authors: Jesse Bricker, Gary V. Engelhardt

Chestnut Hill, Mass.: Center for Retirement Research at Boston College, 2007

MEASUREMENT ERROR IN EARNINGS DATA IN THE HEALTH AND RETIREMENT STUDY

Jesse Bricker and Gary V. Engelhardt*

CRR WP 2007-16
Released: October 2007
Draft Submitted: September 2007

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
Tel: 617-552-1762 Fax: 617-552-0191
http://www.bc.edu/crr

* Jesse Bricker is a graduate student at the Maxwell School of Economics at Syracuse University. Gary V. Engelhardt is an associate professor of economics at Syracuse University and a Center for Policy Research senior research associate. All research with the restricted-access data from the Health and Retirement Study was performed under agreement in the Center for Policy Research at Syracuse University. We thank Dan Black for helpful discussions and comments. All errors are our own. The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium. The opinions and conclusions expressed are solely those of the authors and should not be construed as representing the opinions or policy of SSA, any agency of the Federal Government, or Boston College.

About the Center for Retirement Research

The *Center for Retirement Research at Boston College*, part of a consortium that includes parallel centers at the University of Michigan and the National Bureau of Economic Research, was established in 1998 through a grant from the Social Security Administration. The Center's mission is to produce first-class research and forge a strong link between the academic community and decision makers in the public and private sectors around an issue of critical importance to the nation's future. To achieve this mission, the Center sponsors a wide variety of research projects, transmits new findings to a broad audience, trains new scholars, and broadens access to valuable data sources.

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
phone: 617-552-1762 fax: 617-552-0191
e-mail: crr@bc.edu
www.bc.edu/crr

Affiliated Institutions:
American Enterprise Institute
The Brookings Institution
Center for Strategic and International Studies
Massachusetts Institute of Technology
Syracuse University
Urban Institute

Abstract

We provide new evidence on the extent of measurement error in respondent-reported earnings data by exploiting detailed W-2 records matched to older workers in the Health and Retirement Study (HRS). Our empirical findings are qualitatively consistent with the findings of previous studies. Mean measurement error in the 1991 HRS earnings data for men is somewhat larger than what has been found in other validation studies, but is still modest, averaging about 0.059 log points, approximately 5.9 percent, or $1,500. For women in 1991, it is 0.067 log points, approximately 6.7 percent, or $916. We find a negative correlation between the measurement error and the true value of earnings as measured by the W-2 records, which indicates the presence of non-classical measurement error. For men and women, this error shows little correlation with a standard set of cross-sectional earnings determinants. The one exception is that the measurement error rises with reported education. The bias on the OLS parameter estimate of the impact of having a college degree or higher (relative to a high school drop-out) from using the respondent-reported rather than the W-2 earnings is positive and estimated to be 0.071 log points, or roughly a bias of 7 percent.

I. Introduction

The determinants of earnings play a central role in labor-market studies. As has been long understood in the labor and econometrics literatures, measurement error in respondent-reported earnings in survey data can cause standard econometric estimators, such as the Ordinary Least Squares (OLS) estimator, to generate inefficient, as well as potentially biased and inconsistent, estimates of the determinants of earnings. Although the availability of administrative data sources on earnings has allowed researchers to document the extent of measurement error in important surveys such as the Current Population Survey (CPS), Panel Study of Income Dynamics (PSID), and the Survey of Income and Program Participation (SIPP), little is known about measurement error in earnings in many household surveys in the United States compared to the frequency with which they are used in applied research (Bound, Brown, and Mathiowetz, 2001). Furthermore, what is known has come from validation studies that, in the case of the CPS in 1976-7 (Bound and Krueger, 1991; Bollinger, 1998) and the PSID in 1982-6 (Rodgers, Brown, and Duncan, 1993; Bound, Brown, Duncan, and Rodgers, 1994; Pischke, 1995), are now dated, especially given the many important secular changes in the U.S. labor market in the last three decades.

In this paper, we provide new evidence on the extent of measurement error in respondent-reported earnings data by exploiting detailed earnings data in the Health and Retirement Study (HRS). The HRS is a nationally representative longitudinal survey of the over-50 population. Respondents are interviewed every two years until they die about their income, wealth, health, family structure, housing, and employment. As such, the

HRS is the pre-eminent study of the demography and economics of aging in the United States.

A unique feature of the HRS that we exploit in this analysis is that respondents were asked for their consent to link their survey responses to administrative data on their earnings provided by the Social Security Administration (SSA) and the Internal Revenue Service (IRS). These data include Social Security covered earnings beginning in 1951 and W-2 earnings records beginning in 1978. These earnings are available up through various years from 1991-2003, depending on details of the informed consent agreement, and can be matched to respondent reported earnings from 1991-2003 from the income section of the survey to make estimates of earnings measurement error.

We make a number of contributions to the literature. First, while many previous validation studies (Bound and Krueger, 1991; Bollinger, 1998; Pedace and Bates, 2000) have used Social Security covered earnings records, which are censored at the taxable maximum earnings, we use data from the W-2 earnings records that are free of censoring, and, thus, we sidestep some of data limitations that have complicated previous studies. Second, and more generally, by using data from 1991 and 2003, we provide more up-to-date estimates of the extent of earnings measurement error. Finally, we provide evidence on measurement error for older workers, an increasingly important part of the labor market as the population ages.

Broadly speaking, our empirical findings are qualitatively consistent with the findings of previous studies. Mean measurement error in the 1991 HRS earnings data for men is somewhat larger than what has been found in the CPS, PSID, and SIPP, but is still modest, averaging about 0.059 log points, approximately 5.9 percent, or $1,500. For women in 1991, it is 0.067 log points, approximately 6.7 percent, or $916. We find a negative correlation between the measurement error and the true value of earnings as measured by the W-2 records, which is also similar to previous studies, and indicates the presence of non-classical measurement error. For men and women, this error shows little correlation with a standard set of cross-sectional earnings determinants. The one exception is that the measurement error rises with reported education. The bias on the OLS parameter estimate of the impact of having a college degree or higher (relative to a high school drop-out) from using the respondent-reported rather than the W-2 earnings is positive and estimated to be 0.071 log points, or roughly a bias of 7 percent.

The paper is organized as follows. Section II describes the HRS data, and section III outlines the analytic framework. The empirical results and their relationship to previous findings in the literature are given in section IV. The paper concludes with a summary of the main findings and a discussion of their implications.

II. Data Description

We use detailed data from the Original Cohort of the HRS, those born 1931-41, who entered the study in 1992. Juster and Suzman (1995) give background on the HRS; Moon and Juster (1995) discuss in detail measures of economic status, including income, and the survey design. In the income section of each wave of the survey (given every other year from 1992-2004), one individual in each household—designated as the "financial respondent"—was asked about their own earnings *and* the earnings of the spouse (if present) in the previous calendar year (every other year from 1991-2003, respectively). Specifically, in the first wave, 1992, the previous calendar year was intended to be 1991. Some interviews, however, did not occur until 1993. For those, income was still asked about for 1991. For subsequent waves, the protocol changed so that income was always asked about for the calendar year prior to the year in which the interview occurred.

Some financial respondents either refused to answer the earnings question or answered "don't know." For these cases in the first wave, the respondents were given "range cards" with predetermined values from which to choose their earnings range. In subsequent waves, these respondents were asked a follow-up series of unfolding-bracket questions about earnings. Then, in all waves, the HRS imputed missing earnings by hot-deck methods.[1]

A unique feature of the HRS that we exploit in this study is that respondents were asked for their consent to link their survey responses to administrative data on their earnings provided by the Social Security Administration (SSA) and the Internal Revenue Service (IRS). These data are discussed in detail in Olson (1999) and Mitchell, Olson, and Steinmeier (2000) and include Social Security covered earnings beginning in 1951 and W-2 earnings records beginning in 1978, the latter of which are the basis for our true earnings measure.

The consent process has gone through three main phases. First, in 1992, Original Cohort respondents were asked permission to link to earnings prior to 1992, which resulted in matched Social Security and W-2 records from 1951-91 and 1981-91, respectively. About 75 percent of Original Cohort respondents gave permission in 1992. Those who failed to consent were asked again in 1994 and 1996 for permission to match

[1] This is described in detail in documentation on the HRS website. In particular, see page 10 of http://hrsonline.isr.umich.edu/meta/1996/impute/h1996inf.pdf.

records up through 1991. In 1998, two new cohorts, the War Babies (b. 1942-6) and the Children of the Depression, or CODA, (b. 1924-30), entered the study and the consent form was updated to match earnings up through 1997. All War Babies and CODA members, as well as Original Cohort members who had not yet consented, were asked for permission. Finally, in 2004, another cohort, the Early Baby Boomers (b. 1947-53), entered the study. At this point, the HRS broadened the scope of the consent form to include not only earnings up through 2003, but also to allow for earnings in future years to be added to the matched-earnings files so that respondents would not have to be continually asked permission to update their earnings. All Original Cohort members, including those who already had given permission to match earnings up through either 1991 or 1997, were asked for new permission in 2004. Unfortunately, the HRS does not ask retrospective questions about earnings (in years other than the year prior to the interview year), so that there are no respondent-reported data to be matched to the administrative earnings histories for years prior to 1991.

In this study, we focus on measurement error in earnings for Original Cohort members in 1991 and 2003.[2] In 1991, these individuals were between 50-60 years old; in 2003, they were between 62-72 years old. Table 1 documents how we derived our analysis samples for these years. In column 1 of panel A, there were 5,868 men in the Original Cohort, of which 4,289, or 73 percent, gave consent to match their earnings histories. Of those, 897 were out of the labor force in 1991 and had no W-2 records for that year, leaving 3,392 men, from which we excluded those with IRS Form 1040 Schedule C self-employment income, leaving 2,876 men.

[2] In another paper in progress, we examine the panel data patterns of measurement error in an unbalanced sample of earners drawn from all available HRS cohorts and calendar years.

For our analysis, we use the Medicare earnings from the W-2 records, which are employers' reports to the federal government on earnings on each job in each year, as our measure of true earnings. These records include the value of deferred compensation, such as elective deferrals to 401(k) plans, which are included in the Medicare tax base, but excluded from the federal income tax base (Cunningham and Engelhardt, 2002). Technically, the W-2 records only report earnings up to the Medicare covered-earnings threshold or cap. This cap (in nominal terms) was $125,000 in 1991, $130,200 in 1992, and $135,000 in 1993. Therefore, with the exception of this censoring in these data, the W-2 records give us a complete picture of annual earnings for our sample of HRS workers. Beginning in 1994, the cap was abolished and all earnings were subject to the Medicare payroll tax. As shown in column 1 of the table, there were only 54 men in 1991 with earnings above the Medicare cap, so that any censoring of the data is very minor. Once we exclude those above the Medicare cap and those who did not have positive respondent-reported earnings, we are left we a final analysis sample of 2,670 men in 1991 who had both respondent-reported and W-2 earnings. Similarly, we have a sample of 2,935 women in 1991 (column 2). Only one woman had earnings above the Medicare cap in 1991.

Panel B shows how the 2003 samples were constructed. The main distinctions between 1991 and 2003 were, first, that, by 2003, Original Cohort members were much older (aged 62-72) and far less likely to be in the labor force and have a W-2 record, and second, there was no Medicare cap in 2003, so that all earnings are measured in the W-2s. Overall, there are 635 and 857 men and women, respectively with both respondent-reported and W-2 earnings.

In order to gauge the comparability of the analysis samples to all Original Cohort members in the HRS, Table 2 shows selected sample characteristics for males and females in 1991, respectively. Across a broad array of demographic measures, including age, education, marital status, race, ethnicity, health, and veteran status, there is little difference between the analysis sample members and all Original Cohort members. This is consistent with Haider and Solon (2000), who found that those members of the Original Cohort who gave consent to match administrative earnings did not appear to be a selected group, based on observable characteristics. A similar table for 2003 is available upon request.

III. Analytic Framework

To help frame the empirical analysis that follows, we briefly review the well-known econometric implications of measurement error in cross-sectional models. We focus on dependent-variable measurement error, because earnings are a key outcome in labor-market studies. Specifically, let the true model for a given measure of annual earnings, y, for worker i be

$$y_i = \mathbf{x}_i\beta + \varepsilon_i \,, \tag{1}$$

where $\mathbf{x}$ is a vector of explanatory variables measured without error, and ε is the disturbance term. Assume that $Cov(\mathbf{x}_i, \varepsilon_i) = 0$. Equation (1) represents a prototypical specification in studies of the retirement earnings test, standard Mincerian earnings regressions, and intergenerational earnings mobility specifications, for example.

Let the respondent-reported earnings, y^w, be a combination of true earnings and measurement error,

$$y_i^w = y_i + u_i^w, \tag{2}$$

where u^w is the error. We begin with the simple case of classical measurement error, $Cov(y_i, u_i^w) = 0$, in which the measurement error is not correlated with true earnings. In this standard textbook case, the ordinary least squares (OLS) estimator of β in (1) is consistent, but inefficient, as long as the measurement error is uncorrelated with the explanatory variables in $\mathbf{x}$, i.e., $Cov(\mathbf{x}_i, u_i^w) = 0$.

The more interesting case occurs when the measurement error is non-classical, i.e., when the measurement error is correlated with true earnings, $Cov(y_i, u_i^w) \neq 0$. For example, if this relationship takes the form

$$u_i^w = \delta y_i + v_i^w, \tag{3}$$

in which v^w is white noise, then the earnings equation in (1) written in terms of reported earnings becomes

$$y_i^w = \mathbf{x}_i(1+\delta)\beta + \upsilon_i^w, \tag{4}$$

where $\upsilon^w \equiv (1+\delta)\varepsilon + v^w$. Even if the measurement error is uncorrelated with the explanatory variables in $\mathbf{x}$, (4) illustrates that the OLS estimator is still biased and inconsistent, with proportional bias equal to δ (Bound, Brown, Duncan, and Rodgers, 1994).

With this as background, our empirical analysis is in three parts. First, we take u^w to be the difference between respondent-reported and W-2 earnings and then document basic empirical patterns of measurement error. We consider measurement error that is additive in earnings levels as well as in the log of earnings. The former has been the dependent variable of interest in recent studies of the impact of the Social

Security retirement earnings test on the earnings of older workers (Baker and Benjamin, 1999; Friedberg, 2000; Disney and Smith, 2002; Gruber and Orszag, 2003; Song, 2004; Tran, 2004; Engelhardt and Kumar, 2006; Friedberg and Webb, 2006; Song and Manchester, 2007; Haider and Loughran, forthcoming), some of which has been done using the HRS. The latter is the focal dependent variable in the labor literature and implies that the measurement error is multiplicative in levels.

Second, we estimate the relationship between the measurement error and true earnings, δ, in (3) above and test the null hypothesis of classical measurement error ($\delta = 0$) versus the alternative hypothesis of non-classical measurement error ($\delta \neq 0$), with particular attention to $\delta < 0$, which has been found in the CPS by Bound and Krueger (1992) and the PSIDVS by Bound, Brown, Duncan, and Rodgers (1994). Third, we estimate the relationship between the measurement error and a standard set of explanatory variables found in earnings models, $Cov(\mathbf{x}_i, u_i^w)$, using the following specification,

$$u_i^w = \mathbf{x}_i \alpha + \omega_i^w, \tag{5}$$

where ω is white noise, and test the null hypothesis that the measurement error is unrelated to standard determinants of annual earnings ($\alpha = 0$).

Finally, in recognition that earnings are also an important explanatory variable in empirical studies, we examine the implications of our findings for independent-variable measurement error. Specifically, let the true model for a generic outcome variable, z, be

$$z_i = \phi y_i + \zeta_i, \tag{6}$$

where ζ is a disturbance term. If the measurement error in earnings is non-classical, then the bias in the OLS estimator of (6) when y^w is used in place of y is equivalent to the estimate of γ in the following auxiliary regression,

$$u_i^w = \gamma y_i^w + \xi_i, \tag{7}$$

where ξ is white noise; if the measurement error in earnings is classical, then γ reduces to

$$\frac{\sigma_{u^w}^2}{\sigma_{u^w}^2 + \sigma_y^2}, \tag{8}$$

known as the *variance ratio* (Bound, Brown, Duncan, and Rodgers, 1994). In our analysis, we report estimates of γ and the variance ratio, and compare them to estimates from past studies.

IV. Empirical Results

From the first row of panel A of Table 3, the mean respondent-reported earnings for all men in 1991 are $33,584, with a standard deviation (in parentheses) of $22,733. The mean W-2 earnings are $32,071, with a standard deviation of $20,093. Thus, the mean measurement error, expressed as the difference between the respondent-reported and W-2 earnings, is $1,507, with a standard deviation of $13,899. Figure 1 plots the associated distribution of the measurement error, which is unimodal and symmetric. Overall, 45.1 and 83.3 percent of the errors are within plus or minus $2000 and $10,000, respectively.

Column 6 of Table 3 shows a variance ratio of 0.324, which indicates that 32.4 percent of the variation in the respondent-reported earnings is attributable to the measurement error variance (conversely, 67.6 percent of the variation in the respondent-reported earnings is due to signal, i.e., attributable to variation in true earnings). When earnings are used as an independent variable, and measurement error is classical, then the variance ratio gives the bias in the OLS estimator, which would be 32.4 percent. When the measurement error is non-classical, the bias is given by γ, the estimates of which are shown in column 7 as $\hat{\gamma} = 0.296$ and $\hat{\gamma} = 0.197$ in levels and logs, respectively.

The second row of panel A in Table 3 shows the mean and standard deviation of 1991 earnings for only those men who were financial respondents and, accordingly, reported on their own earnings. Interestingly, this sample restriction, which limits an obvious source of measurement error, does not have an important impact on reducing the mean or variance of the measurement error. The third row includes only men with non-imputed earnings. Not surprisingly, imputation adds a substantial amount of noise, as excluding imputed values reduces the variance ratio by a third. Panel B of Table 3 and Figure 5 show similar statistics about log earnings for men in 1991. Overall, the measurement error patterns for men roughly parallel each other in logs and levels of earnings.

In column 5, we present estimates of the relationship between the measurement error and true earnings, as measured by δ in (3) above, and test the null hypothesis of classical measurement error ($\delta = 0$) versus the alternative hypothesis of non-classical measurement error ($\delta \neq 0$). In panel A, for all men in 1991, $\hat{\delta} = -0.100$, with a standard error (in square brackets) of 0.013. That is for each additional $1000 in true earnings, measurement error falls by $100. We find that the exclusion of imputed values reduces (in absolute value) the estimate of δ to $\hat{\delta} = -0.042$, consistent with Hirsch and

Schumacher (2004), who found that hot-deck procedures can exacerbate "mean reversion" in measurement error. For log earnings in panel B, $\hat{\delta} = -0.304$, with a standard error of 0.012, which can be interpreted as the elasticity of measurement error with respect to true earnings: when true earnings double, measurement error falls by 30 percent. In both levels and logs, the null hypothesis of classical measurement error ($\delta = 0$) can be rejected in favor of the alternative of non-classical measurement error and, in particular, $\delta < 0$.

Qualitatively, these results are similar to what has been found in the CPS by Bound and Krueger (1992) and the PSIDVS by Bound, Brown, Duncan, and Rodgers (1994), referred to as "mean-reverting measurement error." Quantitatively, the estimates of δ and γ for log earnings in panel B are larger than those from the CPS by Bound and Krueger (1992) and the PSIDVS by Bound, Brown, Duncan, and Rodgers (1994), which are reproduced in Table 4, indicating the potential for substantially more bias from measurement error in studies of earnings using the HRS. Whether this is due to differences in sample construction (e.g., previous studies have examined a broader age range of workers than in the HRS) or changes over time in the reporting of labor-market behavior is an open question.

Panels C and D of Table 3 show similar statistics for earning levels and log earnings for men in 2003 (Figures 3 and 7 plot the associated distributions of the measurement error). The results in these panels are mixed. In levels (panel C), the mean error and the variance ratio are substantially larger in 2003 than in 1991, but the estimates of δ are substantially smaller (in absolute value). In fact, the null hypothesis of classical measurement error cannot be rejected at customary significance levels. However, in logs (panel D), while the mean error is larger in 2003 than 1991, the variance ratio is actually smaller, and there is still statistically significant evidence of mean-reverting non-classical measurement error.

Table 5 for women parallels Table 3 for men (Figures 2, 4, 6, and 8 plot the associated distributions of the measurement error). Overall, the mean error for women is less than that for men, similar to Bound and Krueger (1992), but the estimates of δ, the variance ratio, and γ are similar to those for men, both in levels and logs in 1991 and 2003.

Next, we present estimates of α in (5), the relationship between measurement error, u^w, and the correlates of earnings, $\mathbf{x}$, in Table 6. Specifically, $\mathbf{x}$ includes age, dummy variables for education (high school, some college, and college degree or higher; high school drop-out omitted), marital status (married, divorced or separated, and widowed; never married omitted), race (black and other; white omitted), Hispanic ethnicity, self-assessed health status (excellent, very good, good, and fair; poor omitted), veteran status, region of residence, and a full set of one-digit occupation and industry categories. Because respondents are interviewed in different months of the year, we control for the number of months of recall since December of the previous calendar year. Finally, we used the employer identification numbers in the W-2 records to tabulate the number of jobs the respondent held during the year, as the extent of measurement error may depend on the number of jobs the individual had.

Columns 1 and 2 show the OLS estimates for earnings levels for the full sample of men in 1991 and 2003, respectively. Results for the sub-samples with men who were financial respondents and those with non-imputed earnings are available upon request.

The R^2 from the regression in column 1 is 0.03, so that the earnings correlates in **x** jointly explain very little of the variation in measurement error. Qualitatively, the error is statistically significantly higher for men with higher education levels and lower for Hispanic men in 1991. Columns 1 and 2 of Table 7 show the parallel estimates for log earnings. Similar patterns emerge there, with the addition of married men having lower error than never married men in 1991. Finally, columns 3 and 4 of both tables show estimates for women in 1991 and 2003 for earnings levels and log earnings, respectively. Like men, there is little correlation between the measurement error and standard correlates of earnings.

Quantitatively, the estimates in Tables 6 and 7 give the OLS bias from using respondent-reported rather than true earnings as the dependent variable in (1). We illustrate this more clearly in Tables 8 and 9, in which the parameters in (1) are estimated using respondent-reported and true earnings in 1991 for men and women in levels and logs, respectively. For example, in column 1 of Table 9, men with a college degree or higher, earned 49.2 percent more than high school drop-outs, based on respondent-reported earnings, but only 42.1 percent more based on W-2 earnings (in column 2). The difference between the two estimates, 7.1 percent, is what appears in column 1 of Table 7.

V. Summary and Implications

This paper examines the cross-sectional patterns of measurement error in earnings for older workers using detailed data from W-2 records for HRS respondents in 1991 and 2003. Qualitatively, our empirical findings are consistent with those of previous studies.

Mean measurement error in the 1991 HRS earnings data for men is somewhat larger than what has been found in CPS, PSIDVS, and SIPP, but is still modest, averaging about 0.059 log points, approximately 5.9 percent, or $1,500. For women in 1991, it is 0.067 log points, approximately 6.7 percent, or $916. We find a negative correlation between the measurement error and the true value of earnings as measured by the W-2 records, which is also similar to previous studies, and indicates non-classical measurement error. For men and women, this error shows little correlation with a standard set of cross-sectional earnings determinants. The one exception is that the measurement error rises with reported education. The bias on the OLS parameter estimate of the impact of having a college degree or higher (relative to a high school drop-out) from using the respondent-reported rather than the W-2 earnings is positive and estimated to be 0.071 log points, or roughly a bias of 7 percent.

An important area of recent research on the annual earnings of older workers has been on the impact of the Social Security retirement earnings test (Baker and Benjamin, 1999; Friedberg, 2000; Disney and Smith, 2002; Gruber and Orszag, 2003; Song, 2004; Tran, 2004; Engelhardt and Kumar, 2006; Friedberg and Webb, 2006; Song and Manchester, 2007; Haider and Loughran, forthcoming). Before 2000, the earnings test reduced benefits received for Social Security beneficiaries over age 65 who earned more than a minimum threshold; beginning in 2000, the test was abolished. While some studies of the 2000 abolition of the earnings test have used SSA administrative earnings data, such as Song (2004) and Song and Manchester (2007), others, such as Engelhardt and Kumar (2006) and Friedberg and Webb (2006) have used HRS respondent-reported data. The fact that, in the current study, the measurement error in reported earnings does not seem to differ significantly according to age or calendar year suggests that studies of the impact of the retirement earnings test that exploit age and time variation in HRS respondent-reported earnings probably do not suffer from first-order bias from measurement error in earnings beyond that induced by the non-classical measurement error.[3]

Although we limited the analysis to cross-sectional models, we hope to complement this paper with future research examining four areas. First, we are currently exploiting the SSA covered earnings and W-2 records to examine measurement error in respondent-reported labor force participation using both earnings receipt information from the income section of the HRS survey, as well as responses to questions on working for pay from the employment section of the survey. Second, we also have combined W-2 and respondent-reported earnings across multiple years from 1991-2003 to make an unbalanced panel of earnings with which to analyze measurement error in panel earnings models. In particular, in panel data, the attenuation from measurement error is a function of the serial correlation of the measurement error and true earnings, respectively. The panel we have assembled allows us to estimate the autocorrelation in measurement error over a longer time horizon than has been used in previous studies, e.g., Bound, Brown, Duncan, and Rodgers (1994), Pischke (1996), and Bound and Krueger (1992). Third, we have begun to examine measurement error in HRS earnings in a framework that treats the W-2 earnings themselves as potentially noisy measures of true earnings (Black, Berger, Scott, 1999; Barron, Berger, Black, 2000; Kane, Rouse, and Staiger, 1999; Kapteyn and

[3] Interestingly, Haider and Loughran (forthcoming), using respondent-reported and administrative data from the CPS, actually find that respondent-reported data understate the amount of bunching of earnings at the earnings test threshold and, therefore, understate the impact of the earnings test on labor supply in structural models.

Ypma, 2007). Finally, we hope to use the HRS to examine the impact of measurement error on estimates of earnings dynamics and inequality (Baker, 1997; Haider, 2001; Baker and Solon, 2003). In particular, there are a number of studies in that literature and the macroeconomics literature on consumption inequality and precautionary saving (Carroll and Samwick, 1998; Blundell and Preston, 1998, among others) that estimate the variance in permanent and transitory components of annual earnings from respondent-reported earnings in major household surveys, such as the PSID. The key question is to what extent do the estimates of the variance in, say, transitory earnings reflect a true transitory component versus measurement error in reported earnings. The rich data we have developed for the HRS can be used to examine this issue.

References

Baker, Michael, "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics* 15 (1997): 338-375.

Baker, Michael, and Dwayne Benjamin, "How Do Retirement Tests Affect the Labour Supply of Older Men?" *Journal of Public Economics* 71 (1999): 27-52.

Baker, Michael, and Gary Solon, "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Tax Records," *Journal of Labor Economics* 21 (2003): 289-321.

Barron, John M., Mark C. Berger, and Dan A. Black, "The Use of Matched Employee-Employer Data as a Means of Assessing Data Reliability with Nonclassical Measurement Error," in J.C. Haltiwanger, J.I. Lane, J.R. Spletzer, J.J.M. Theeuwes, and K.R. Troske (eds.) *The Creation and Analysis of Employer-Employee Matched Data* (Amsterdam: Elsevier), 1999, pp. 419-438.

Black, Dan A., Mark C. Berger, and Frank A. Scott, "Bounding Parameter Estimates with Nonclassical Measurement Error," *Journal of the American Statistical Association* 95 (2000): 739-748.

Blundell, Richard, and Ian Preston, "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics* 106 (1998): 603-640.

Bollinger, Christopher R., "Measurement Error in the Current Population Survey," *Journal of Labor Economics*, 16 (1998): 576-594.

Bound, John, Charles Brown, Gregory J. Duncan, and Willard L. Rodgers, "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data," *Journal of Labor Economics* 12 (1994): 345-368.

Bound, John, Charles Brown, and Nancy Mathiowetz, "Measurement Error in Survey Data," in J.J. Heckman and E. Leamer (eds.) *Handbook of Econometrics, Volume 5* (Amsterdam: North-Holland) 2001, pp. 3705-3843.

Bound, John, and Alan Krueger, "The Extent of Measurement Error in Longitudinal Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics*, 9 (1991): 1-24.

Carroll, Christopher D., and Andrew A. Samwick, "How Important is Precautionary Saving?" *Review of Economics and Statistics* 80 (1998): 410-419.

Cunningham, Christopher R., and Gary V. Engelhardt. 2002. "Federal Tax Policy, Employer Matching, and 401(k) Saving: Evidence from HRS W-2 Records," *National Tax Journal*, 55: 617-645.

Disney, Richard, and Sarah Tanner, “The Labour Supply Effect of the Abolition of the Earnings Rule for Older Workers in the United Kingdom,” *Economic Journal* 112 (2002): 136-152.

Duncan, Gregory, and Daniel Hill, “An Investigation of the Extent and Consequences of Measurement Error in Labor Economic Survey Data,” *Journal of Labor Economics*, 3 (1985): 508-522.

Engelhardt, Gary V., and Anil Kumar, “The Repeal of the Retirement Earnings Test and the Labor Supply and Health of Older Men,” Mimeo., Syracuse University, 2006.

Friedberg, Leora, “The Labor Supply Effects of the Social Security Earnings Test,” *Review of Economics and Statistics* 82:1 (2000): 48-63.

Friedberg, Leora, and Anthony Webb, “Persistence in Labor Supply and the Response to the Social Security Earnings Test,” Mimeo., University of Virginia, 2006.

Griliches, Zvi, and Jerry A. Hausman, “Errors in Variables in Panel Data,” *Journal of Econometrics* 31 (1986): 93-118.

Gruber, Jonathan, and Peter Orszag, “Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?” *National Tax Journal* 56:4 (2003): 755-773.

Haider, Steven J., “Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991,” *Journal of Labor Economics* 19 (2001): 799-836.

Haider, Steven J., and David S. Loughran, “Do the Elderly Respond to Taxes on Earnings? Evidence from the Social Security Earnings Test,” *Journal of Human Resources*, forthcoming.

Haider, Steven J., and Gary Solon, “Nonrandom Selection in the HRS Social Security Earnings Sample,” Working Paper No. 00-01, RAND Labor and Population Program, 2000.

Haider, Steven J., and Gary Solon, “Life-cycle Variation in the Association between Current and Lifetime Earnings,” *American Economic Review*, 2006, forthcoming.

Hirsch, Barry T., and Edward J. Schumacher, “Match Bias in Wage Gap Estimates Due to Earnings Imputations,” *Journal of Labor Economics* 22 (2004): 689-722,

Juster, F. Thomas, and Richard Suzman, “An Overview of the Health and Retirement Study.” *Journal of Human Resources* 30 (1995): S7-S56.

Kane, Thomas, Cecilia E. Rouse, and Douglas O. Staiger, “Estimating Returns to Schooling when Schooling is Misreported,” NBER Working Paper Bo. 7235, 1999.

Kapteyn, Arie, and Yelmer Ypma, "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data," *Journal of Labor Economics* 25 (2007): 513-551.

Mitchell, Olivia S., Jan Olson, and Thomas L. Steinmeier, "Social Security Earnings and Projected Benefits," in Olivia S. Mitchell, P. Brett Hammond, and Anna M. Rappaport, eds., *Forecasting Retirement Needs and Retirement Wealth* (Philadelphia: University of Pennsylvania Press), 2000, pp. 327-360.

Moon, Marilyn, and F. Thomas Juster, "Economic Status Measures in the Health and Retirement Study," *Journal of Human Resources* 30 (1995): S138-S157.

Olson, Janice A., "Linkages with Data from Social Security Administrative Records in the Health and Retirement Study." *Social Security Bulletin* 62 (1999): 73-85.

Pedace, Roberto, and Nancy Bates, "Using Administrative Records to Assess Earnings Reporting Error in the Survey of Income and Program Participation," *Journal of Economic and Social Measurement* 26 (2000): 173-192.

Pischke, Jorn-Steffen, "Measurement Error and Earnings Dynamics: Some Estimates for the PSID Validation Study," *Journal of Business and Economic Statistics* 13 (1995): 305-314.

Rodgers, Willard, Charles Brown, Gregory Duncan, "Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages," *Journal of the American Statistical Association* 88 (1993): 1208-1218.

Song, Jae G., "Evaluating the Initial Impact of Eliminating the Retirement Earnings Test," *Social Security Bulletin* 65:1 (2004): 1-15.

Song, Jae G., and Joyce Manchester, "New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000," *Journal of Public Economics* 91 (2007): 669-700.

Tran, Bac, "The Effect of the Repeal of the Retirement Earnings Test on the Labor Supply of Older Workers." Mimeo., University of Maryland, 2004.

Table 1. Construction of the Sample for the HRS Original Cohort, by Year of Earnings

Sample	(1) Men	(2) Women
A. 1991 Earnings		
Number in Cohort	5868	6784
Number of Consents to Match 1991 Earnings Records	4289	5097
Number with W-2 Earnings in 1991	3392	3412
Number with No Self-Employment Income	2876	3124
Number below Medicare Cap	2822	3123
Number with Respondent-Reported Earnings	2670	2935
B. 2003 Earnings		
Number in Cohort	4065	5297
Number of Consents to Match 2003 Earnings Records	2392	2837
Number with W-2 Earnings in 2003	1030	1167
Number with No Self-Employment Income	818	1023
Number with Respondent-Reported Earnings	635	857

Note: The Medicare cap was $125,000 in 1991; there was no cap in 2003.

Table 2. Means of Selected Characteristics of Analysis Sample with Matched W-2 Records to All Workers in the HRS for 1991 Earnings, by Sex

	(1)	(2)	(3)	(4)
	Men		Women	
	Sample	All	Sample	All
Age	55.879	56.897	53.171	53.700
	(4.560)	(5.240)	(5.373)	(5.447)
High School Drop-out	0.233	0.271	0.189	0.265
High School Degree	0.374	0.344	0.422	0.401
Some College	0.194	0.187	0.216	0.195
College Degree or Higher	0.200	0.199	0.173	0.139
Married	0.885	0.872	0.742	0.758
Never Married	0.023	0.029	0.033	0.029
Divorced	0.081	0.085	0.150	0.133
Widowed	0.011	0.013	0.076	0.081
White	0.770	0.741	0.745	0.704
Black	0.130	0.148	0.170	0.177
Other Race	0.019	0.022	0.021	0.023
Hispanic	0.081	0.089	0.065	0.095
Excellent Health	0.239	0.218	0.266	0.227
Very Good Health	0.304	0.269	0.321	0.281
Good Health	0.309	0.292	0.277	0.270
Fair Health	0.109	0.137	0.109	0.147
Poor Health	0.039	0.084	0.027	0.076
Veteran	0.564	0.562	0.010	0.008
Number of Recall Months	8.682	8.846	8.605	8.754
	(2.801)	(2.820)	(2.768)	(2.846)
Number of Jobs	1.270		1.268	
	(0.638)		(0.582)	
N	2670	5868	2935	6784

Note: Standard deviations in parentheses. The number of jobs held is based on employer identification numbers found in the W-2 records and, therefore, is not available for the all HRS individuals in columns 2 and 4, respectively.

Table 3. Summary Statistics for Earnings and Measurement Error between Self-Reported and W-2 data for Men in 1991 and 2003

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Means (Standard Deviations)					
Sample	N	Respondent-Reported Earnings	W-2 Earnings	Error	δ	Variance Ratio	γ
A. Earnings Levels in 1991							
All	2670	33,584	32,071	1507	-0.100	0.324	0.296
		(22,733)	(20,093)	(13,899)	[0.013]		[0.010]
Financial Respondents Only	1885	35,926	33,837	2089	-0.090	0.341	0.320
		(24,089)	(20,827)	(14,988)	[0.016]		[0.012]
Non-Imputed Only	2421	33,821	32,345	1476	-0.042	0.205	0.183
		(21,683)	(20,023)	(10,158)	[0.010]		[0.009]
B. Log Earnings in 1991							
All	2670	10.155	10.096	0.059	-0.304	0.322	0.197
		(0.857)	(0.921)	(0.634)	[0.012]		[0.014]
Financial Respondents Only	1885	10.224	10.149	0.075	-0.287	0.305	0.175
		(0.866)	(0.932)	(0.617)	[0.014]		[0.016]
Non-Imputed Only	2421	10.169	10.113	0.057	-0.236	0.259	0.130
		(0.848)	(0.904)	(0.535)	[0.011]		[0.013]

Table 3. (Continued) Summary Statistics for Earnings and Measurement Error between Self-Reported and W-2 data for Men in 1991 and 2003

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Means (Standard Deviations)					
Sample	N	Respondent-Reported Earnings	W-2 Earnings	Error	δ	Variance Ratio	γ
			C. Earnings in 2003				
All	635	35,985	31,496	4489	-0.008	0.496	0.496
		(43,095)	(30,731)	(30,470)	[0.039]		[0.020]
Financial Respondents Only	457	37,520	32,799	4721	-0.061	0.510	0.510
		(45,421)	(32,797)	(33,439)	[0.048]		[0.025]
Non-Imputed Only	556	36,368	32,981	3387	0.014	0.248	0.254
		(37,194)	(31,908)	(18,336)	[0.024]		[0.018]
			D. Log Earnings in 2003				
All	635	9.922	9.832	0.089	-0.173	0.158	0.204
		(1.215)	(1.192)	(0.739)	[0.024]		[0.023]
Financial Respondents Only	457	9.940	9.343	0.112	-0.207	0.171	0.185
		(1.247)	(1.733)	(0.787)	[0.028]		[0.028]
Non-Imputed Only	556	9.974	9.867	0.107	-0.190	0.241	0.136
		(1.183)	(1.222)	(0.688)	[0.023]		[0.024]

Note: Standard deviations in parentheses. Standard errors in square brackets.

Table 4. Findings from Previous Studies for Men

	(1)	(2)	(3)
Year and Earnings Measure	δ	Variance Ratio	γ
A. CPS			
1977 Log Earnings	-0.197	0.181	0.026
1976 Log Earnings	-0.194	0.156	-0.016
B. PSIDVS			
1986 Log Earnings	-0.172	0.302	0.239
1982 Log Earnings	-0.104	0.151	0.076

Note: Panel *A* originates from Bound and Krueger (1991), but is taken from Table 2 in Bound, Brown, Duncan and Rodgers (1994). Panel *B* is taken from Table 1 in Bound, Brown, Duncan and Rodgers (1994).

Table 5. Summary Statistics for Earnings and Measurement Error between Self-Reported and W-2 data for Women in 1991 and 2003

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Means (Standard Deviations)					
Sample	N	Respondent-Reported Earnings	W-2 Earnings	Error	δ	Variance Ratio	γ
			A. Earnings Levels in 1991				
All	2935	19,694	18,790	916	-0.142	0.349	0.320
		(15,426)	(13,853)	(10,146)	[0.013]		[0.011]
Financial Respondents Only	1625	20,573	19,439	1134	-0.131	0.357	0.328
		(15,982)	(14,049)	(10,475)	[0.018]		[0.014]
Non-Imputed Only	2631	19,398	18,952	446	-0.096	0.282	0.247
		(15,159)	(13,840)	(8673)	[0.012]		[0.010]
			B. Log Earnings in 1991				
All	2935	9.569	9.503	0.067	-0.289	0.305	0.174
		(0.916)	(0.987)	(0.654)	[0.011]		[0.013]
Financial Respondents	1625	9.619	9.551	0.069	-0.265	0.301	0.184
		(0.909)	(0.957)	(0.628)	[0.015]		[0.017]
Non-Imputed Only	2631	9.562	9.515	0.048	-0.239	0.237	0.085
		(0.898)	(0.984)	(0.548)	[0.010]		[0.012]

Table 5. (Continued) Summary Statistics for Earnings and Measurement Error between Self-Reported and W-2 data for Women in 1991 and 2003

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Means (Standard Deviations)					
Sample	*N*	Respondent-Reported Earnings	W-2 Earnings	Error	δ	Variance Ratio	γ
			C. Earnings in 2003				
All	857	25,411	23,895	1517	-0.052	0.194	0.165
		(23,928)	(22,454)	(10,998)	[0.017]		[0.015]
Financial Respondents Only	519	25,125	23,688	1436	-0.055	0.211	0.184
		(22,962)	(21,337)	(11,044)	[0.023]		[0.020]
Non-Imputed Only	728	25,949	24,759	1190	-0.031	0.156	0.137
		(24,600)	(23,213)	(9979)	[0.016]		[0.014]
			D. Log Earnings in 2003				
All	857	9.651	9.327	0.031	-0.126	0.184	0.211
		(1.179)	(1.413)	(0.672)	[0.020]		[0.018]
Financial Respondents Only	519	9.623	9.327	0.010	-0.110	0.195	0.234
		(1.218)	(1.419)	(0.699)	[0.027]		[0.023]
Non-Imputed Only	728	9.681	9.641	0.040	-0.113	0.195	0.127
		(1.158)	(1.149)	(0.566)	[0.018]		[0.017]

Note: Standard deviations in parentheses. Standard errors in square brackets.

Table 6. OLS Parameter Estimates of the Determinants of Measurement Error in Earnings Levels, by Sex, in 1991 and 2003, Respectively, with Standard Errors in Parentheses

	(1)	(2)	(3)	(4)
	Men		Women	
	Error in 1991 Earnings	Error in 2003 Earnings	Error in 1991 Earnings	Error in 2003 Earnings
Age	11	400*	-64*	-42
	(63)	(241)	(37)	(64)
High School Degree	248	-302	454	-81
	(796)	(3898)	(563)	(1239)
Some College	1520*	216	672	1877*
	(797)	(3525)	(519)	(1022)
College Degree or Higher	2578***	11,132***	1262**	1301
	(865)	(3,540)	(599)	(1134)
Married	-812	3178	953	1667
	(1847)	(9271)	(1095)	(2258)
Divorced	-239	-113	593	2491
	(2071)	(10,287)	(1175)	(2412)
Widowed	-1887	312.	976	347
	(3299)	(11070)	(1268)	(2421)
Black	-210	-1936	-693	1660
	(897)	(4119)	(552)	(1212)
Other Race	-2635	-5592	-1417	-1696
	(2065)	(7790)	(1349)	(2342)
Hispanic	-2599**	-2489	-725	-108
	(1105)	(4568)	(829)	(1552)
Excellent Health	1060	-2711	380	2417
	(1530)	(7878)	(1248)	(2900)
Very Good Health	-677	-1379	373	3029
	(1502)	(7446)	(1233)	(2794)
Good Health	-871	1063	-10	3043
	(1488)	(7353)	(1229)	(2806)
Fair Health	-673	-2371	870	2008
	(1632)	(7649)	(1299)	(2901)
Veteran	-758	1391	669	-1237
	(607)	(2692)	(1944)	(3411)
Number of Recall Months	-20	-4187	230	915
	(486)	(2604)	(319)	(796)
Number of Jobs	346	2521	-208	-232
	(444)	(2035)	(332)	(665)
N	2532	610	2857	838
R^2	0.03	0.05	0.01	0.04

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level. The specifications also include a full set of dummy variables for one-digit occupation, industry, as well as region of residence.

Table 7. OLS Parameter Estimates of the Determinants of Measurement Error in Log Earnings, by Sex, in 1991 and 2003, Respectively, with Standard Errors in Parentheses

	(1)	(2)	(3)	(4)
	Men		Women	
	Error in 1991 Earnings	Error in 2003 Earnings	Error in 1991 Earnings	Error in 2003 Earnings
Age	-0.003	0.007	-0.005**	0.003
	(0.003)	(0.006)	(0.002)	(0.004)
High School Degree	0.044	-0.090	0.032	-0.075
	(0.036)	(0.096)	(0.035)	(0.076)
Some College	0.059	-0.001	0.042	0.029
	(0.036)	(0.087)	(0.033)	(0.063)
College Degree or Higher	0.071*	0.112	0.073*	0.051
	(0.039)	(0.087)	(0.038)	(0.070)
Married	-0.183**	0.087	0.099	0.098
	(0.084)	(0.228)	(0.069)	(0.139)
Divorced	-0.118	0.018	0.080	0.049
	(0.094)	(0.254)	(0.074)	(0.149)
Widowed	-0.244	0.025	0.129	0.034
	(0.150)	(0.273)	(0.080)	(0.149)
Black	-0.021	0.022	-0.076**	0.085
	(0.041)	(0.102)	(0.035)	(0.075)
Other Race	-0.048	-0.149	-0.071	-0.056
	(0.094)	(0.192)	(0.085)	(0.144)
Hispanic	-0.095*	-0.023	-0.090*	-0.054
	(0.050)	(0.113)	(0.052)	(0.096)
Excellent Health	0.034	0.083	0.064	0.311*
	(0.070)	(0.194)	(0.078)	(0.179)
Very Good Health	-0.007	-0.047	0.059	0.294*
	(0.068)	(0.183)	(0.077)	(0.172)
Good Health	-0.031	-0.082	0.080	0.285*
	(0.068)	(0.181)	(0.077)	(0.173)
Fair Health	-0.041	-0.230	0.069	0.198
	(0.074)	(0.188)	(0.082)	(0.179)
Veteran	0.005	0.050	0.033	-0.056
	(0.028)	(0.066)	(0.122)	(0.210)
Number of Recall Months	0.015	0.012	0.033	0.006
	(0.022)	(0.064)	(0.020)	(0.049)
Number of Jobs	0.006	0.147***	-0.037*	0.0001
	(0.020)	(0.050)	(0.021)	(0.0410)
N	2532	610	2857	838
R^2	0.02	0.07	0.02	0.04

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level. The specifications also include a full set of dummy variables for one-digit occupation, industry, as well as region of residence.

Table 8. A Comparison of the OLS Parameter Estimates of the Determinants of Earnings Levels Using the Respondent-Reported and W-2 Record Earnings, by Sex, in 1991, with Standard Errors in Parentheses

	(1)	(2)	(3)	(4)
	Men		Women	
	Respondent-Reported Earnings	W-2 Earnings	Respondent-Reported Earnings	W-2 Earnings
Age	-328***	-338***	-76	-13
	(90)	(77)	(51)	(43)
High School Degree	-5014***	-5262***	-1593**	-2047***
	(1133)	(985)	(764)	(649)
Some College	5499***	3979***	5554***	4882***
	(1135)	(986)	(704)	(598)
College Degree or Higher	18,671***	16,093***	14,346***	13,084***
	(1232)	(1070)	(812)	(690)
Married	12,624***	13,436***	-5138***	-6092***
	(2630)	(2284)	(1484)	(1262)
Divorced	9569***	9808***	-3120*	-3713***
	(2948)	(2561)	(1593)	(1355)
Widowed	5124	7011*	-4169**	-5145***
	(4696)	(4079)	(1720)	(1462)
Black	-3496***	-3286***	192	885
	(1278)	(1110)	(748)	(636)
Other Race	-6497**	-3862	107	1524
	(2940)	(2554)	(1829)	(1555)
Hispanic	-7474***	-4876***	-3052***	-2327**
	(1574)	(1367)	(1124)	(956)
Excellent Health	9870***	8810***	6297***	5917***
	(2179)	(1893)	(1692)	(1438)
Very Good Health	7224***	7900***	5078***	4704***
	(2138)	(1857)	(1671)	(1421)
Good Health	4366**	5236***	4095**	4105***
	(2119)	(1841)	(1666)	(1416)
Fair Health	2750	3423*	3986**	3116**
	(2323)	(2018)	(1761)	(1498)
Veteran	-1295	-537	4617*	3949*
	(865)	(751)	(2636)	(2241)
Number of Recall Months	442	462	72	-158
	(692)	(601)	(433)	(368)
Number of Jobs	-3504***	-3850***	-2066***	-1858***
	(632)	(549)	(450)	(383)
N	2532	2532	2857	2857
R^2	0.26	0.28	0.21	0.27

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level. The specifications also include a full set of dummy variables for one-digit occupation, industry, as well as region of residence.

Table 9. A Comparison of the OLS Parameter Estimates of the Determinants of Log Earnings Using the Respondent-Reported and W-2 Record Earnings, by Sex, in 1991, with Standard Errors in Parentheses

	(1)	(2)	(3)	(4)
	Men		Women	
	Respondent-Reported Earnings	W-2 Earnings	Respondent-Reported Earnings	W-2 Earnings
Age	-0.025***	-0.022***	-0.006*	0.0001
	(0.003)	(0.004)	(0.003)	(0.0030)
High School Degree	-0.164***	-0.208***	-0.144***	-0.176***
	(0.043)	(0.048)	(0.045)	(0.050)
Some College	0.160***	0.101**	0.265***	0.223***
	(0.043)	(0.048)	(0.042)	(0.046)
College Degree or Higher	0.492***	0.421***	0.627***	0.554***
	(0.047)	(0.052)	(0.048)	(0.053)
Married	0.499***	0.683***	-0.208**	-0.307***
	(0.101)	(0.110)	(0.088)	(0.096)
Divorced	0.391***	0.510***	-0.023	-0.103
	(0.113)	(0.124)	(0.095)	(0.104)
Widowed	0.002	0.246	-0.079	-0.208*
	(0.180)	(0.197)	(0.102)	(0.112)
Black	-0.077	-0.057	0.045	0.121**
	(0.049)	(0.054)	(0.045)	(0.049)
Other Race	-0.142	-0.094	-0.011	0.060
	(0.113)	(0.124)	(0.109)	(0.119)
Hispanic	-0.275***	-0.181***	-0.225***	-0.135*
	(0.060)	(0.066)	(0.067)	(0.073)
Excellent Health	0.462***	0.428***	0.579***	0.515***
	(0.083)	(0.092)	(0.101)	(0.110)
Very Good Health	0.420***	0.427***	0.519***	0.460***
	(0.082)	(0.090)	(0.099)	(0.109)
Good Health	0.305***	0.336***	0.497***	0.417***
	(0.081)	(0.089)	(0.099)	(0.108)
Fair Health	0.174*	0.216**	0.374***	0.304***
	(0.089)	(0.098)	(0.105)	(0.115)
Veteran	0.018	0.014	0.222	0.189
	(0.033)	(0.036)	(0.157)	(0.171)
Number of Recall Months	-0.175***	-0.181***	-0.154***	-0.117***
	(0.024)	(0.027)	(0.027)	(0.029)
Number of Jobs	-0.025***	-0.022***	-0.006*	0.0001
	(0.003)	(0.004)	(0.003)	(0.0030)
N	2532	2532	2857	2857
R^2	0.22	0.19	0.18	0.17

Note: * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level. The specifications also include a full set of dummy variables for one-digit occupation, industry, as well as region of residence.

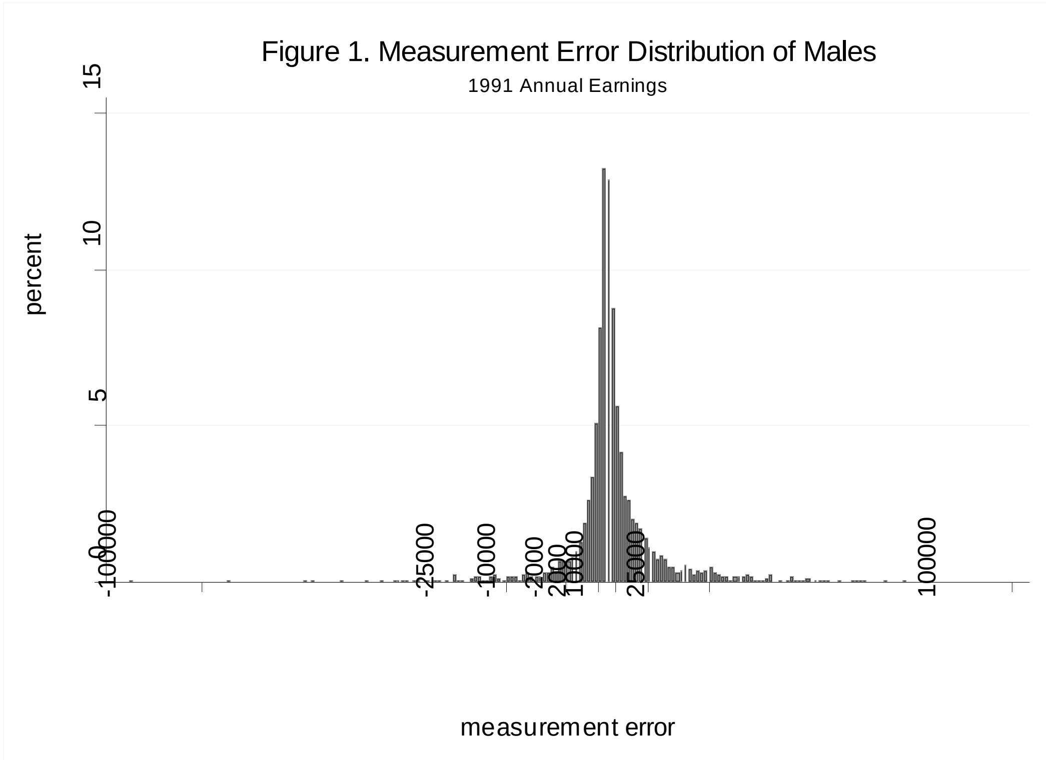

Figure 1. Measurement Error Distribution of Males
1991 Annual Earnings
percent
15
10
5
0
-100000
-25000
-10000
-2000
2000
10000
25000
100000
measurement error

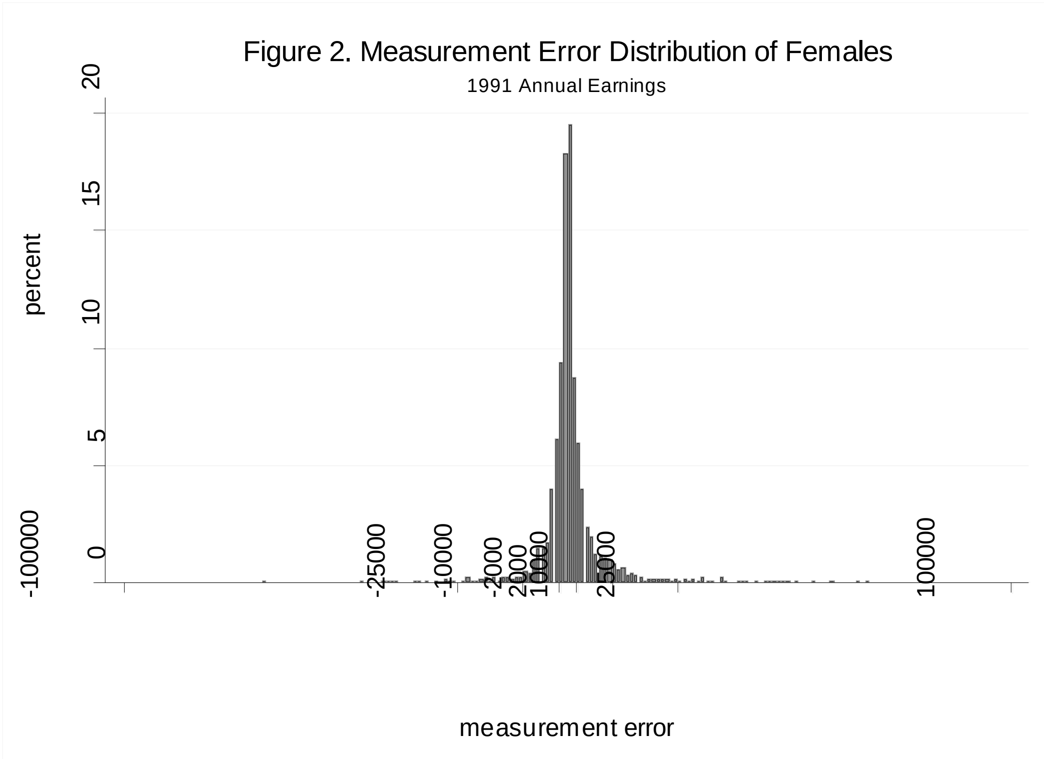

Figure 2. Measurement Error Distribution of Females
1991 Annual Earnings
percent
20
15
10
5
0
-100000
-25000
-10000
-2000
2000
10000
25000
100000
measurement error

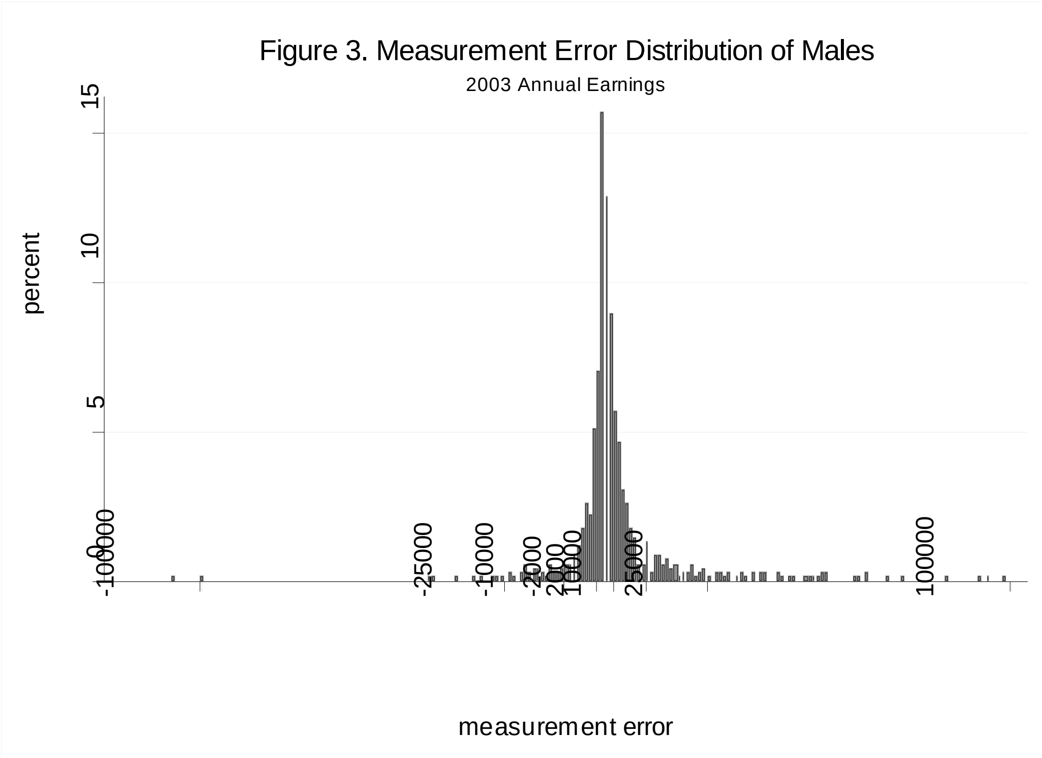
Figure 3. Measurement Error Distribution of Males
2003 Annual Earnings
percent
0
5
10
15
-100000
-25000
-10000
-2000
2000
10000
25000
100000
measurement error

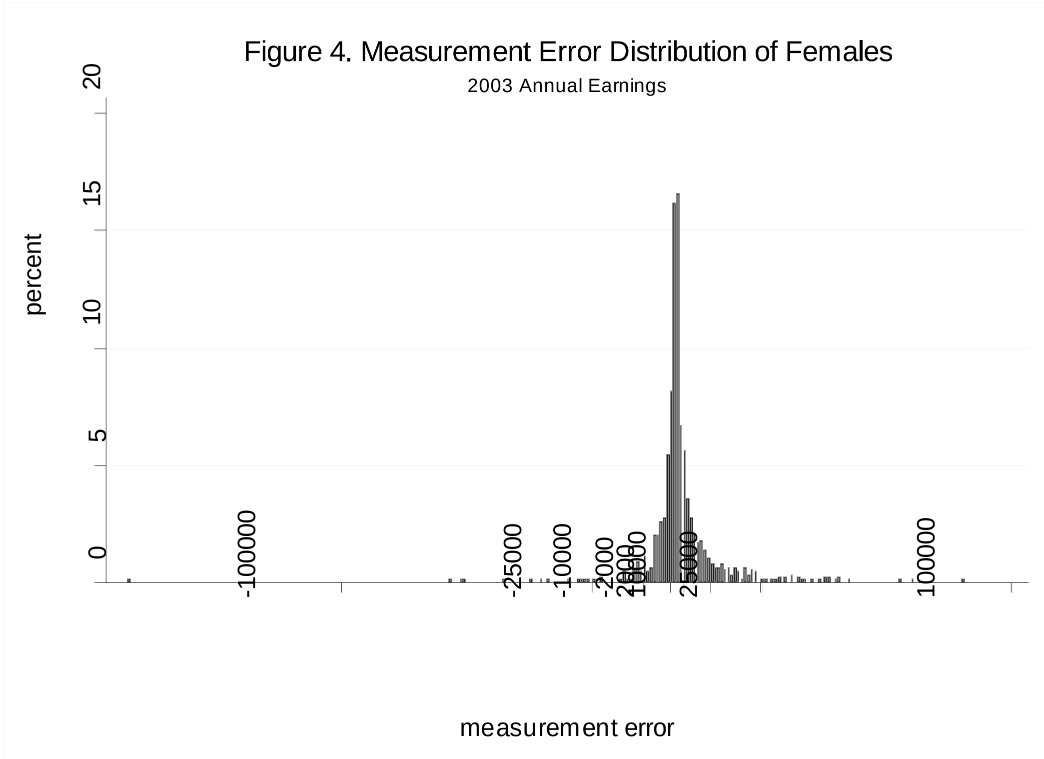
Figure 4. Measurement Error Distribution of Females
2003 Annual Earnings
percent
0
5
10
15
20
-100000
-25000
-10000
-2000
2000
10000
25000
100000
measurement error

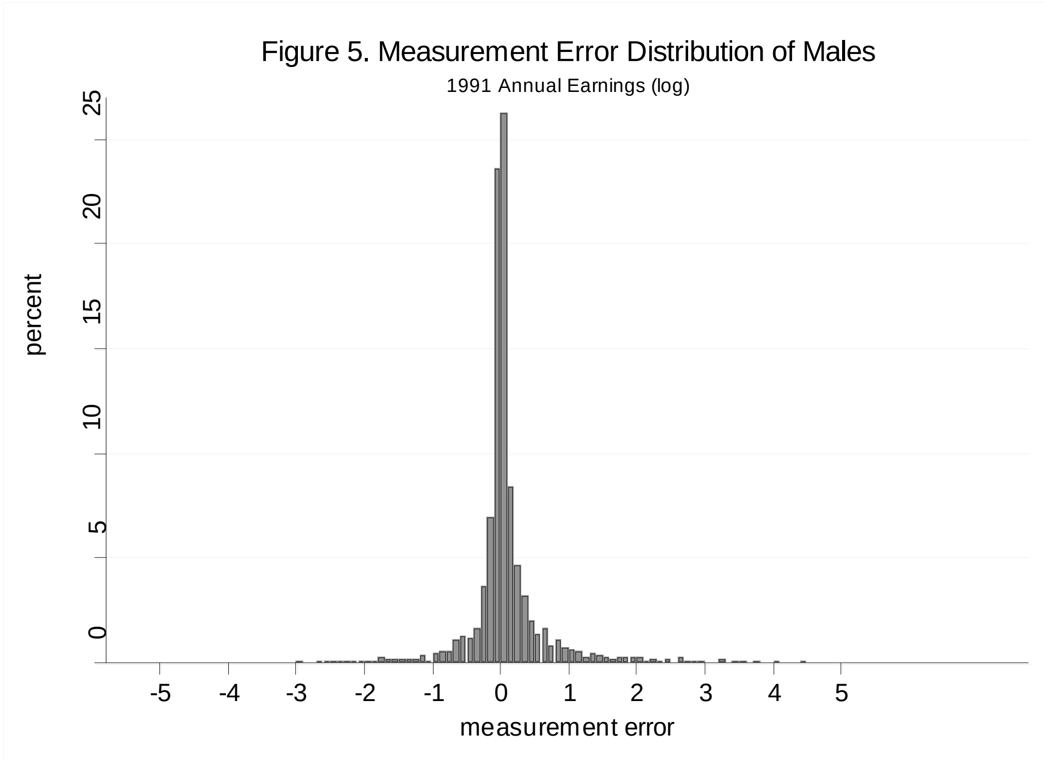
Figure 5. Measurement Error Distribution of Males
1991 Annual Earnings (log)
percent
0
5
10
15
20
25
-5
-4
-3
-2
-1
0
1
2
3
4
5
measurement error

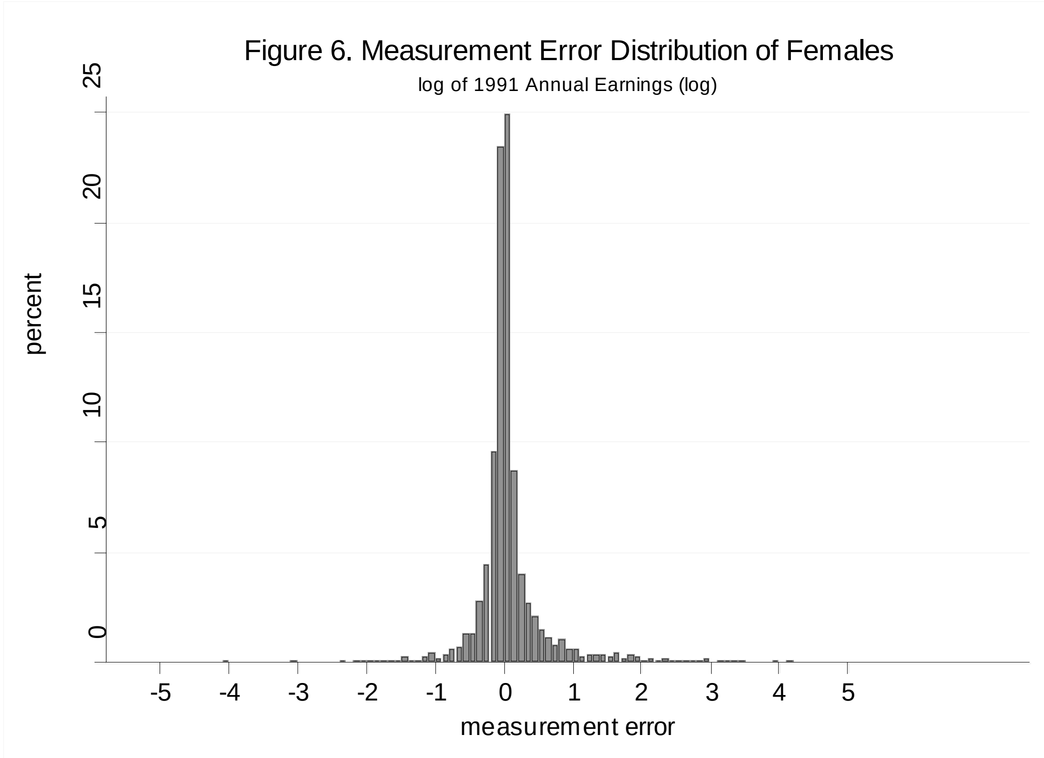
Figure 6. Measurement Error Distribution of Females
log of 1991 Annual Earnings (log)
percent
0
5
10
15
20
25
-5
-4
-3
-2
-1
0
1
2
3
4
5
measurement error

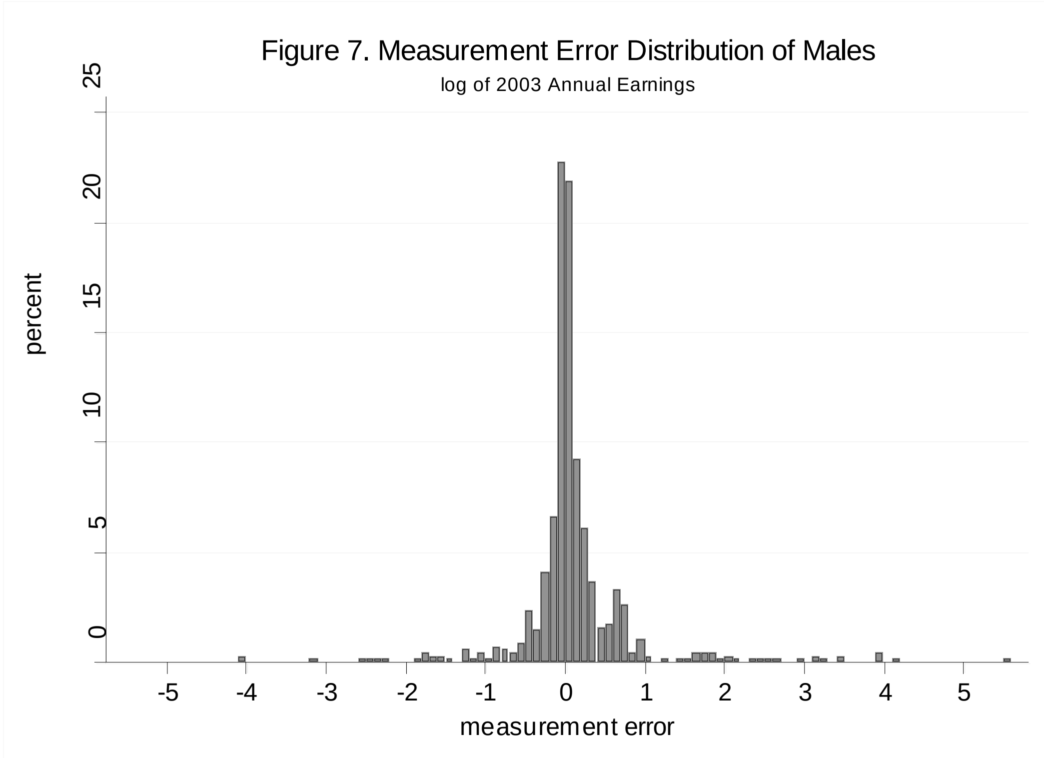
Figure 7. Measurement Error Distribution of Males
log of 2003 Annual Earnings
percent
0
5
10
15
20
25
-5
-4
-3
-2
-1
0
1
2
3
4
5
measurement error

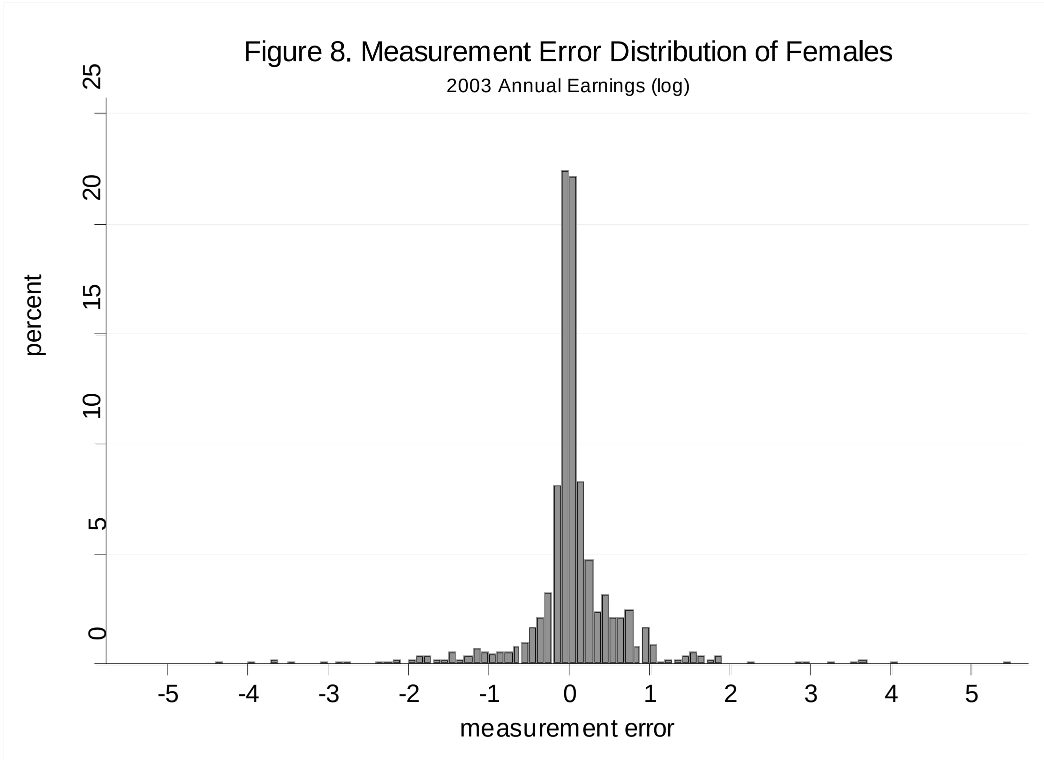

Figure 8. Measurement Error Distribution of Females
2003 Annual Earnings (log)
percent
0
5
10
15
20
25
-5
-4
-3
-2
-1
0
1
2
3
4
5
measurement error

RECENT WORKING PAPERS FROM THE CENTER FOR RETIREMENT RESEARCH AT BOSTON COLLEGE

Evaluating the Advanced Life Deferred Annuity – An Annuity People Might Actually Buy
Guan Gong and Anthony Webb, September 2007

Population Aging, Labor Demand, and the Structure of Wages
Margarita Sapozhnikov and Robert K. Triest, September 2007

Work at Older Ages: Is Raising the Early Retirement Age an Option for Social Security Reform?
John A. Turner, August 2007

The Labor Supply of Older Americans
Alicia H. Munnell and Steven A. Sass, May 2007

Why Do Japanese Workers Remain in the Labor Force So Long?
John B. Williamson and Masa Higo, May 2007

Literacy, Trust and the 401(k) Savings Behavior
Julie R. Agnew, Lisa Szykman, Stephen P. Utkus, and Jean A. Young, May 2007

The Recent Evolution of Pension Funds in the Netherlands: the Trend to Hybrid DB-DC Plans and Beyond
Eduard H.M. Ponds and Bart van Riel, May 2007

Demographic Influences on Saving-Investment Balances in Developing and Developed Economies
Ralph C. Bryant, May 2007

Social Security Spouse and Survivor Benefits for the Modern Family
Melissa M. Favreault and C. Eugene Steuerle, February 2007

How Economic Security Changes During Retirement
Barbara Butrica, February 2007

International Investment for Retirement Savers: Historical Evidence on Risk and Returns
Gary Burtless, February 2007

All working papers are available on the Center for Retirement Research website (http://www.bc.edu/crr) and can be requested by e-mail (crr@bc.edu) or phone (617-552-1762).